

29. 05. 00

EV

1800 / 189

**ROMANIA**

REC'D 31 MAY 2000

WIPO

PCT

**OFICIUL DE STAT PENTRU INVENȚII ȘI MĂRCI**

Nr.1006540/03.05.2000

**CERTIFICAT DE PRIORITATE**

Nr.007/03.05.2000

**PRIORITY  
DOCUMENT**SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Certificăm că descrierea anexată este copia identică a descrierii  
invenției cu titlul:

"RECUNOASTEREA VORBIRII SI ANALIZA SEMNALELOR PRIN CAUTAREA  
DIRECTA A SUBSECVENTELOR CU MASURI MAXIME DE INCREDERE"

pentru care s-a constituit depozitul reglementar al cererii de  
brevet de invenție la Oficiul de Stat pentru Invenții și Mărci,  
la data de .....25.02.1999..... sub nr. ....99-00214.....  
de către S.C. "FIGHTER" SERV S.R.L., Baia Mare, RO

D I R E C T O R   G E N E R A L



18

OFICIUL DE STAT PENTRU INVENȚII ȘI MĂRCI
Cerere de brevet de invenție
Nr. 99 - 00214
Data depozit 25.02.99

## Recunoașterea Vorbirii și Analiza Semnalelor prin căutarea directă a subsecvențelor cu Măsurî Maxime de Încredere

### 5. 1 Domeniile tehnice în care invenția poate fi utilizată sint

- Recunoașterea Vorbirii,
- Recunoașterea Cuvintelor Cheie,
- Analiza și Recunoașterea Imaginilor,
- 10. • Recunoașterea Segmentelor În Studiul Materiilor Organice (Studiul Genomului)

### 2 Stadiul tehnicii

- 15. • La recunoașterea cuvintelor cheie se utilizează adesea reprezentarea segmentelor nedorite printr-un model atotcuprinzător de umplutură <Garbage/Filler Models>. Se utilizează modele fonetice dependente/independente de context și fără constrângeri lexicale. Se efectuează apoi o estimare <VITERBI> (probabilitatea maximă a unei căi), procedura înainte-înapoi (probabilitatea considerând toate căile) sau o reducere a modelului la Programare Dinamică. O alternativă permite relaxarea constrângerii asupra poziționării punctelor de început și de sfârșit ale cuvintelor în secvențele rostite [3]. Dar aceasta necesită o normalizare ulterioară.
- 20. • Una din cele mai performante metode actuale realizează recunoașterea cuvintelor cheie prin modelarea segmentelor nedorite utilizând media celor mai performante  $K$  corespondente în esanțion. [2]
- 25. • Studiul genomului se face, în anumite laboratoare (ISREC), pe baza reprezentării secvențelor distanțate de proteine cu funcționalități comune sub forma unor modele numite 'Profile Generalizate'. Cu ajutorul acestora se caută subsecvențele corespondente de DNA. Algoritmul este

14

5. bazat pe Programarea Dinamica cu modificari minore datorate specificitatii modelului. Deoarece calculele necesare sint extrem de costisitoare, sisteme paralele de pina la 16 calculatoare (INSECT) sint utilizate. In aceste zile, o implementare hardware bazata pe FPGA-uri a fost realizata [4]. Aceasta calculeaza in jur de 7 milioane celule esantion-stare pe secunda.

- 10.
- Pentru recunoasterea obiectelor in imaginii exista o gama larga de metode. Se utilizeaza cross-corelatia, 'snakes', analiza trasaturilor, a texturii, a culorii, etc. Metodele au avantaje si deavantaje diferite pentru fiecare aplicatie. In domeniul cartografiei se utilizeaza adesea 'snakes' ce trebuie initializati de catre operator. In domeniul roboticii se utilizeaza, printre altele, analiza conturilor si cross-corelatia proiectiilor.

### 3 Expunerea Inventiei

15. Descriem inventia utilizind notatii si termeni legati de recunoasterea vorbirii si utilizind probabilitati posterioare. Metodele ramn identice daca se folosesc probabilitati sau alte tipuri de aproximare a acestora bazate pe aplicatia corespunzatoare, ca de exemplu anumite distante intre valorile culorii pixelilor. Scorurile utilizate in profilele secventelor biomoleculare sint de asemenea acceptabile pentru estimari de posteriori.

20. Deasemenea noi folosim minusul logaritmului posteriorilor. Metoda este desigur identica daca se folosesc logaritmi pozitivi cu schimbarea semnului comparatiilor. Acestea sind detalii de implementare.

#### 3.1 Descrierea teoretica a problemei

25. In Recunoasterea Vorbirii s-a recomandat [1] utilizarea de masuri de confidenta (incredere) pentru reestimarea primelor  $N$  ipoteze. Modele propuse contin normalizarea probabilitatii posterioare cumulate peste lungimea secventei recunoscute. O alta masura este dubla normalizare peste numarul de foni si numarul de esantioane acustice in fiecare fon. Aceste masuri nu pot fi calculate direct utilizind Viterbi deoarece factorul de normalizare este diferit pentru cai de lungimi diferite.

30. Fie  $X = \{x_1, x_2, \dots, x_n, \dots, x_N\}$  secventa esantioanelor acustice in care cautam subsecventa  $W$  reprezentata ca un Model Markov Ascuns (HMM)

5.

inainte  $M$  avind starile:  $Q = \{q_1, q_2, \dots, q_\ell, \dots, q_L\}$ . In spiritul consideratiilor asupra masurilor de incredere cu normalizare simpla, potrivind  $M$  cu subsecventa  $X_b^e = \{x_b, \dots, x_e\}$  si utilizind starea de umplutura  $q_G$  inainte si dupa  $M$ , obtinind un HMM  $M'$ , putem calcula logaritmul posterior al modelului  $M$ , in modul urmator:

$$\begin{aligned} -\log P(M|X_b^e) &\simeq \frac{1}{e-b+1} \min_{\forall \bar{Q} \in M} -\log P(\bar{Q}|X_b^e) \\ &\simeq \frac{1}{e-b+1} \min_{\forall \bar{Q} \in M} \{-\log P(q^b|q_G) - \sum_{n=b}^{e-1} [\log P(q^n|x_n) \\ &\quad + \log P(q^{n+1}|q^n)] - \log P(q^e|x_e) - \log P(q_G|q^e)\} \quad (1) \end{aligned}$$

10.

unde  $\bar{Q} = \{q^b, q^{b+1}, \dots, q^e\}$  reprezinta una din caile de lungime  $e-b+1$  in  $M$ , iar  $q^n$  este starea HMM vizitata la momentul  $n$  de-a lungul lui  $\bar{Q}$ .  $q_G$  reprezinta modelul de umplutura.

Pentru Recunoasterea Cuvintelor Cheie, definim scorul de recunoastere ca:

$$S(M|X) = \min_{\forall \{b,e\}} -\log P(M|X_b^e) \quad (2)$$

15.

si notam cu:

$$<\bar{S}', in, out> = \operatorname{argmin}_{\{\bar{Q}, b, e\}} \frac{1}{e-b+1} -\log P(\bar{Q}|X_b^e) \quad (3)$$

Cu  $Q = \{\overbrace{q_G, \dots, q_G}^{b-1}, q^b, q^{b+1}, \dots, q^e, \overbrace{q_G, \dots, q_G}^{N-e}\}$ , notam calea ce extinde  $\bar{Q}$  cu  $b-1$  stari  $q_G$  inaintea lui  $q^b$  si  $N-e$  stari  $q_G$  urmind  $q^e$ . In cazul utilizarii starilor de umplutura, calea  $S'$  este de obicei estimata:

20.

$$\begin{aligned} S' &\simeq \operatorname{argmin}_{\forall Q \in M'} -\log P(Q|X) \\ &\simeq \operatorname{argmin}_{\forall Q \in M'} \{-\sum_{n=1}^{b-1} \log P(q_G|x_n) - \log P(q^b|q_G) \\ &\quad - \sum_{n=b}^{e-1} [\log P(q^n|x_n) + \log P(q^{n+1}|q^n)] - \log P(q^e|x_e) \\ &\quad - \log P(q_G|q^e) - \sum_{n=e+1}^N \log P(q_G|x_n)\} \quad (4) \end{aligned}$$

- pentru care  $P(q_G|x_n)$  este egala cu o constanta  $\varepsilon$  sau cu o variabila ca in cazurile mentionate in stadiul tehnicii. Acest procedeu nu conducea in general la solutia definita in (2) si (3). In spiritul normalizarii simple noi propunem metoda Programarii Dinamice Iterative, descrisa in continuare,
5. pentru calcularea unei valori pentru  $\varepsilon$  ce garanteaza obtinerea segmentarii 'optime' definite in (2). Acelasi rezultat ar putea fi obtinut prin normalizarea posteriorului cumulat pentru toate perechile posibile de puncte de inceput si sfirsit, si ar necesita un numar de teste proportional cu patratul lungimii secventei in care se cauta. Metoda inventata are calitatea aditionala de Timp
10. Real la normalizarea simpla.
- Vom nota cu  $S_a^b$  calea S in HMM  $M'$  daca ea intra la esantionul  $b$  in starea  $q_1$  din starea  $q_G$  si iese la esantionul  $e$  din starea  $q_L$  spre starea  $q_G$ .
- Daca un obiect necesita pentru reprezentarea sa HMM tranzitii inapoi, se va realiza reprezentarea sa utilizind mai multe modele HMM inainte si care
15. vor fi tratate ca subsecvente diferite.

### 3.2 Metoda cu Normalizare Simpla

Metoda iterativa de calculare a lui  $\varepsilon$ :

1.  $\varepsilon$  se initializeaza cu *Valoarea\_De\_Initializare*
2.  $\langle w', S \rangle = \text{Viterbi}(\varepsilon)$
20. 3. ciclu:  
 $S' = S$   
 $\varepsilon = w'$
4.  $\langle w', S \rangle = \text{ViterbiI}(\varepsilon)$
5. daca  $S'$  diferit de  $S$   
 mergi la pasul 3 (ciclu)  
 altfel  
 raporteaza raspunsul  $S'$  retestat cu masuri aditionale de incredere
- 25.
- Unde submetoda  $\text{ViterbiI}$  consta in:  
 $\text{ViterbiI}(\varepsilon)$
30. 1. Calculeaza calea  $S_e^\varepsilon$  folosind Viterbi clasic al  $M'$  peste  $X$ .

$$2. w = -\log P(M|X_b^e)$$

3. Raporteaza  $\langle w, S \rangle$

5. Valoarea de initializare a lui  $\epsilon$  se poate alege la intimplare fara pierderea convergentei. Noi am obtinut rezultate foarte bune initializind-ul cu valoarea logaritmului posteriorului obtinut cu metodele descrise in starea tehnicii. In toate testele, aceasta metoda a convers in maximum 4 iteratii.

Demonstratiile corectitudinii vor fi publicate in scurt timp in literatura internationala de specialitate.

10. Metoda este de citeva ordine de marime mai rapida decit alte abordari care ar garanta optimul teoretic definit de masura de incredere cu normalizare simpla. In plus, ea nu necesita un spatiu in memoria de calcul mai mare decit nici una din metodele cunoscute. Rezultatele in recunoasterea cuvintelor cheie se compara pozitiv cu rezultatele tehnicilor ce le cunoastem.

### 3.3 Metoda cu Normalizare Dubla

15. Masura de incredere corespunzatoare se defineste ca:

$$\frac{1}{Nr.Foni\ Parcursi} \sum_{fon_i \in Fonii\ Parcursi} \frac{\sum_{fon_i} - \log posteriori}{lungime\ fon_i} \quad (5)$$

Aceasta metoda consta intr-un algoritm de tip 'breadth first Beam search' caracterizat de absenta modelului de umplutura, un set complex de reguli de reducere si de efectuarea anumitor normalizari:

20. Starea  $q_G$  are in aceasta metoda logaritmul posteriorului de emisie 0. Pentru fiecare esantion  $e$  si pentru fiecare stare  $s$  se calculeaza setul de probabilitati/cai de a avea esantionul  $e$  in starea  $s$ , ca primele  $N$  maxime ale masurii de confidanta pentru caile in HHM  $M'$  de lungime  $e$  si care se termina cu starea  $s$ . Se pot elimina, deasemenea, cele care, potrivit regulilor de reducere, vor pierde cu siguranta cursa finala in favoarea unor cai deja cunoscute.

25. Notam cu  $a_1, p_1, l_1, a_2, p_2$  si  $l_2$  masura de incredere pe fonii precedenti, posteriorul in fonul curent si lungimea in fonul curent pentru calea  $Q_1$  respectiv  $Q_2$ . Regulele ce pot fi utilizate la reducerea spatiului cailor, 'pruning', pentru eliminarea lui  $Q_1$  in favoarea lui  $Q_2$  pot fi in acest caz fiecare din urmatoarele:

$$1. l_2 \geq l_1, A > 0, B \leq 0 \text{ si } L_c^2 A + L_c B + C \geq 0$$

2.  $l_2 \geq l_1$ ,  $A \geq 0$ ,  $B \geq 0$  si  $C \geq 0$

3.  $l_2 \geq l_1$ ,  $A \leq 0$ ,  $C \geq 0$  si  $L^2 A + LB + C \geq 0$

4.  $l_2 \geq l_1$ ,  $A = 0$ ,  $B < 0$  si  $LB + C \geq 0$

5. unde  $A = a_1 - a_2$ ,  $B = (a_1 - a_2)(l_1 + l_2) + p_1 - p_2$ ,  $C = (a_1 - a_2)l_1 l_2 + p_1 l_2 - p_2 l_1$ ,  
 $L = L_{max} - \max\{l_1, l_2\}$ ,  $L_c = -B/2A \geq 0$  iar  $L_{max}$  este lungimea maxima acceptata pentru un fon - echivalentul lungimii acceptate a unei stari din abordarile clasice.

10. Eliminarea unei cai doar daca una din aceste reguli e satisfacuta, garantteaza optimul definit de masura de incredere cu normalizare dubla, daca nici un fon nu poate fi evitat de HMM-ul  $M$  utilizat. Fiecare HMM inainte poate fi descompus in HMM-uri cu aceasta calitate. Regula a 4-a este inclusa in regula a 3-a si testarea sa e inutila daca aceasta din urma a fost deja testata.

15. Primul test,  $l_2 \geq l_1$  ne spune de fapt daca  $Q_2$  are sanse sa elimine  $Q_1$ , in caz contrar se va verifica daca  $Q_1$  elimina  $Q_2$ . Aceste teste se deduc din conditiile asigurarii masurii de incredere maxime dupa efectuarea reducerilor. Pentru a utiliza metoda dublei normalizari fara a descompune HMM-uri ce evita anumiti foni, regulile anterioare se modifica considerind numarul de foni anteriori pentru fiecare cale  $F_1$  respectiv  $F_2$  si numarul de foni ce pot urma starii curente.

20. Un test simplificat poate fi:

- $l_2 \geq l_1$ ,  $A \geq 0$ ,  $p_1 \geq p_2$  respectiv  $F_2 \geq F_1$  in cazul HHM-urilor ce pot evita foni.

25. Acest test este mai slab decit regula 2. Exemplu: O cale este eliminata de o alta daca prima are o masura de confidenta inferioara (numeric mai mare) pe fonii anteriori, o lungime mai mica si minusul logaritmului posteriorului cumulat in fonul curent deasemenea inferior (numeric mai mare) celei de a doua. O masura aditionala de incredere bazata pe lungimea maxima,  $L_{max}$ , si pe maximul minusului logaritmului posteriorului cumulat si normalizat in fon,  $P_{max}$ , e utila pentru garantarea unei limitari puternice a numarului de cai stocate.

30.

- $p > L_{max} P_{max}$  intr-o stare oarecare
- $\frac{p}{l} > P_{max}$  la iesirea din fon

unde  $p$  si  $l$  sint valorile in fonul curent pentru minusul logaritmului posteriorului cumulat si pentru lungimea caii ce se elimina. Aceste teste permit eliminarea cailor prea lungi si care nu sint deosebit de performante, respectiv a cailor cu foni avind un scor inacceptabil, compensat de scoruri foarte bune in alti foni.

5.

Daca  $N$  se alege ca unu, regulile mentionate nu mai sint necesare, ci intotdeauna se propaga calea cu estimarea curenta a masurii de incredere maxime. Rezultatele obtinute sint deosebit de bune, chiar daca optimul definit se garanteaza pentru aceasta metoda doar cind  $N$  este mai mare ca lungimea secventei permise de  $L_{max}$  sau a secventei ce se testeaza.

10.

Aceeasi abordare e functionala pentru normalizarea simpla unde HMM-ul subsecventei va fi format dintr-un singur fon.

### 3.4 Metoda Potrivirii Reale

Noi am definit, de asemenea, o noua masura de incredere care reprezinta bine cerintele recunoasterii. Deoarece fonii si starile absente pot fi modelate de catre HMM-urile utilizate, este normal sa se ceara potrivirea fiecarui fon din model cu o portiune din secventa. Prin urmare, noi consideram masura de incredere a unei subsecvente ca fiind egala cu maximul pentru toti fonii, a minusului logaritmului posteriorului cumulat al fonului, normalizat cu lungimea sa.

15.

20.

$$\max_{fon \in Foni \text{ Parcursi}} \frac{\sum_{fon} - \log posteriori}{lungime fon} \quad (6)$$

Regula ce poate fi utilizata in acest cadru la reducerea numarului de cai explorate este:

25.

- $Q_2$  se elimina in favoarea altei cai  $Q_1$  daca masura de incredere a potrivirii reale pentru fonii anteriori este inferioara (numeric superioara) pentru  $Q_2$  fata de  $Q_1$ , si daca  $p_1 \leq p_2$  si  $l_2 \leq l_1$

unde  $p_1$ ,  $l_1$ ,  $p_2$ ,  $l_2$  reprezinta minusul logaritmului posteriorului cumulat respectiv lungimea masurata in esantioane in fonul curent pentru calea  $Q_1$  respectiv  $Q_2$ .

30.

Asemenea metodei precedente, limitarea numarului de cai considerate se face prin eliminarea celor care corespund criteriilor:

- $p > L_{max} P_{max}$  intr-o stare oarecare



- $\frac{p}{l} > P_{max}$  la iesirea din fon

unde  $p$  si  $l$  sint valorile in fonul curent pentru minusul logaritmului posteriorului cumulat si pentru lungimea caii ce se elimina. Amintim ca semnificatiile constantelor sint lungimea maxima,  $L_{max}$ , si respectiv maximul acceptat al minusului logaritmului posteriorului cumulat si normalizat in fon,  $P_{max}$ .

5.

### 3.5 Performante

Metodele prezentate anterior au fost testate in forma descrisa la punctul 4.2. Implementarea a testat aplicarea acestora pentru recunoasterea cuvintelor cheie. S-a construit curba ROC si s-a constatat ca metodele au dat rezultate superioare celor publicate in literatura de specialitate pentru alte experiente. Metoda dublei normalizari a dat rezultate superioare metodei simplei normalizari.

10.

## 4 Obiectul Inventiei

### 4.1 Obiectul inventiei consta in:

15.

- metoda de recunoastere a subsecventelor utilizind calculul direct de masuri de incredere.
- masura de incredere si metoda de recunoastere numita 'potrivire reala' si bazata pe potrivirea separata a fiecarui fon.
- metodele de recunoastere utilizind normalizarea simpla si dubla

20.

- insotirea acestor masuri cu masurile aditionale de incredere mentionate aici, respectiv normalizarea dubla, lungimea maxima si limitarea potrivirii reale

25.

- Utilizarea metodelor anterioare la recunoasterea cuvintelor cheie
- Utilizarea metodelor anterioare la recunoasterea subsecventelor in structura materiilor organice
- Utilizarea metodelor anterioare la recunoasterea de obiecte in imagini

## 4.2 Mod de executare, folosire si functionare

Executarea: Este necesar un sistem de calcul, dar metoda poate de asemenea fi utilizata intr-o implementare hardware.

- 5. • Se va obtine o reprezentare sub forma HMM a subsecventelor cautate (cuvint, profil proteine, sectiune imagine obiect).
- Se obtine un instrument (eventual antrenat Ex: pentru recunoasterea vorbirii) de estimare a posteriorilor. De exemplu multi-gaussiani, retele neuronale, clusteri, baze de date cu Profile Generalizate, etc.
- Se implementeaza unul din algoritmi propusi.
- 10. Pentru primul algoritm
  - Viterbi se implementeaza utilizind algoritmul clasic Viterbi cu modificarea ca, pentru fiecare pereche  $P = \langle \text{esantion}, \text{stare} \rangle$  se propaga si momentele tranzitiilor intre starea  $q_G$  si starile HMM  $M$  a cãii ce ajunge la  $P$ . Acestea se mostenesc de la calea ce cistiga intrarea in perechea  $P$ , cu exceptia momentului cind se face decizia lor si anume cind ele primesc valoarea esantionului corespunzator.
  - 15. –  $w = -\log P(M|X_b^e)$  se calculeaza scazind din posteriorul cumulat calculat de algoritmul Viterbi pentru calea  $Q_b^e$  valoarea  $(N - (e - b + 1)) * \epsilon$  corespunzind umpluturii si divizind rezultatul prin  $e - b + 1$ .
  - 20. Desigur, factorul  $e - b + 1$  din formula precedenta va fi scos in afara fractiei.
  - Initializarea lui  $\epsilon$  se face cu o valoare medie asteptata. Se poate folosi si  $w$  calculat cind starea  $q_G$  este asociata cu un posterior de emisie egal cu media celor mai bune  $K$  probabilitati de emisie ale esantionului curent. In acest caz,  $K$  se antreneaza utilizind tehnicile actuale cunoscute.
  - 25.

30. Algoritmi urmatori, de tip 'Beam search', se implementeaza dupa descrierea din sectiunea corespunzatoare. Pentru fiecare pereche  $P = \langle \text{esantion}, \text{stare} \rangle$  se calculeaza pentru fiecare cale asociata suma si lungimea din ultimul fon, precum si suma peste cumulumul normalizat al fonilor precedenti (si numarul lor). Deasemenea, esantioanele de intrare si iesire in/din HMM-ul  $M$  sint calculate si propagate ca in metoda precedenta, pentru a asigura localizarea subsecventei.

- 4
- Dacă una și aceeași subsecvență poate avea mai multe modele HMM, toate acestea se iau în considerare ca competitori. Acesta e cazul cuvintelor cu multiple moduri de pronunție sau obiectelor ce pot fi văzute diferit în diferite stări, pentru recunoașterea acestora în imagini.
5. După calculul măsurii de încredere pentru fiecare model de subsecvență, se elimină cele cu o măsură de confidență în dezacord cu un 'threshold' antrenat pentru configurația și scopul aplicației. De exemplu, pentru recunoașterea vorbirii cu rețele neuronale și minus al logaritmului posteriorilor se va alege un 'threshold' maximal în punctul de funcționare dorit al curbei ROC obținută în teste.
10. Alternativ, ramasele se extrag în ordinea măsurii de încredere și cu eliminarea alternativelor în conflict până la epuizare. De fiecare dată când o alternativă e eliminată, subsecvența caută cu HMM corespunzător e reestimată pentru porțiunea rămasă disponibilă ei în secvența în care se operează căutarea. Dacă noua măsură de încredere promovează testul 'threshold'-ului, atunci ea va fi inserată în poziția corespunzătoare valorii în coada de alternative.
15. Alternativ, cu succes pot suporta teste de nivel superior ca de exemplu o întrebare de confirmare pentru recunoașterea vorbirii, opinia unui operator, etc.
20. Pentru recunoașterea obiectelor în imagini:  
Posteriorii se obțin prin calculul unei distanțe între culoarea stării modelului și cea a elementului din secțiunea imaginii. Dacă condițiile de utilizare o impun, imaginea va fi preprocesată pentru a asigura o anumită normalizare (Ex: condiții schimbătoare de lumină vor face necesară o transformare de corecție bazată pe histogramă).
25. Fonilor din recunoașterea vorbirii le corespund părți ale obiectului. Structura (existența tranzițiilor și probabilitatea lor) poate fi schimbată în funcție de caracteristicile detectate ale căii curente. De exemplu, detectând zone ale obiectului cu anumite lungimi, putem estima lungimea așteptată a zonelor ramase. Se va stabili astfel numărul de esanțioane așteptat pentru stările viitoare și se va configura corespunzător HMM-ul atasat obiectului.
30. Se va scana o direcție pentru detectarea unei potriviri optime și ulterior se vor scana celelalte direcții pentru a descoperii noi potriviri, cit și
- 35

8

pentru a verifica cele existente. Testul final va fi certificat prin cross-corelatie sau prin analiza contururilor in pozitia generata de ipoteze.

## 5 Aplicare Industriala

Aici sint prezentate citeva exemple ale aplicatiilor metodei propuse in industrie:

- 5. • Recunoasterea Cuvintelor Cheie incepe sa fie utilizata in raspunatoare automate ale sistemelor bancare, telefonice precum si in automate de control, vinzare sau informare. Metoda ofera o posibilitate de recunoastere a cuvintelor cheie in vorbire continua cu vorbitori multipli.
- 10. • Recunoasterea secventelor in ADN este importanta in studiul genomului uman. Una din cele mai mari probleme ale tehnicilor implicate consta in marea cantitate de date ce trebuie prelucrate.
- 15. • Recunoasterea Obiectelor in Imagini isi are utilizarea, printre altele, in cartografie si in conducerea robotilor industriali. Metoda permite estimarea rapida a pozitiei obiectelor in scena si poate fi urmata de teste suplimentare utilizind metode clasice de cross-corelatie.

## References

- 19. [1] G. Bernardis and H. Bourlard. Improving posterior based confidence measures in hybrid hmm/ann speech recognition systems. In *Proceedings to ICSLP'98 (Sydney, Australia)*, 1998.
- 20. [2] H. Bourlard, B. D'hoore, and J.-M. Boite. Optimizing recognition and rejection performance in wordspotting systems. In *Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (Adelaide, Australia)*, pages I:373-376, 1994.
- 25. [3] J.S. Bridle. An efficient elastic-template method for detecting given words in running speech. In *Brit. Acoust. Soc. Meeting*, pages 1-4, 73.

- [4] E. Mosanya. *A Reconfigurable Processor for Biomolecular Sequence Processing*. PhD thesis, Swiss Federal Institute of Technology EPFL, LSL, February 1999.

# Recunoașterea Vorbirii și Analiza Semnalelor prin căutarea directă a subsecvențelor cu Măsuri Maxime de Încredere

Revendicarea Independenta 1.

5.

Preambul:

Realizeaza recunoasterea unor subsecvente reprezentate ca Modele Markov Ascunse(HMM), cautate intr-o secventa oarecare. Ne referim la masuri de incredere utilizate la reclasarea ipotezelor cistigatoare din Recunoasterea Vorbirii. Exemple de astfel de masuri sint:

10.

normalizarea simpla = posteriorul cumulat normalizat cu lungimea subsecventei

15.

normalizarea dubla = dubla normalizare peste numarul de foni si peste numarul de esantioane acustice in fiecare fon al posteriorului cumulat.

caracterizata prin aceea ca: Permite masura suplimentara a increderii bazate pe extremele valorilor logaritmului posteriorului cumulat al fiecarui fon, normalizat cu lungimea sa. Aceasta masura am denumit-o 'potrivirea reala'.

20.

$$\max_{fon \in Foni \text{ Parcursi}} \frac{\sum_{fon} - \log posteriori}{lungime fon}$$

caracterizata prin aceea ca: Se cauta subsecventa ce ofera maximizarea uneia dintre masurile de incredere amintite, peste toate potrivirile posibile.

25.

caracterizata prin aceea ca: Permite reevaluarea alternativelor ce ofera maximizarea unei masuri de incredere mentionata, pe baza altei masuri de incredere.

29.

caracterizata prin aceea ca: Calculeaza alternativa ce maximizeaza 'normalizarea simpla' utilizind metoda ce noi am numit-o 'Programare Dinamica Iterativa' si care estimeaza posteriorul de emisie al starilor de umplutura, in mod iterativ, ca fiind egal cu masura de incredere in iteratia anterioara.

5. caracterizata prin aceea ca: Calculeaza alternativa ce maximizeaza 'normalizarea simpla', 'normalizarea dubla' sau 'potrivirea reala' prin utilizarea unui algoritm ce considera posteriorul de emisie al starii de umplutura ca nul. Aceasta metoda calculeaza progresiv, pentru fiecare pereche de esantion si stare a HMM, un set de alternative posibile de a ajunge acolo.

Calcularea acestui set se bazeaza pe seturile de alternative ce conduc la starile ce pot fi asociate esantionului anterior celui curent.

Acest set poate fi redus utilizind reguli caracteristice masurii de incredere si care asigura corectitudinea inferentei.

10. Acest set poate deasemenea fi redus utilizind heuristici bazate pe regulile mentionate, pentru a mari viteza de calcul cu riscul diminuarii calitatii teoretice a recunoasterii.

9

# Recunoașterea Vorbirii și Analiza Semnalelor prin căutarea directă a subsecvențelor cu Măsuri Maxime de Încredere

Revendicarea Dependentă 2.

5.

Preambul:

Bazata pe Revendicarea 1.

Realizeaza estimarea existentei cuvintelor cheie si a pozitiei lor in secvente rostite.

10.

caracterizata prin aceea ca: Utilizeaza metodele descrise in Revendicarea 1. pentru recunoasterea subsecvențelor reprezentate prin modele markov ascunse.



3

5

# Recunoașterea Vorbirii și Analiza Semnalelor prin căutarea directă a subsecvențelor cu Măsuri Maxime de Încredere

Revendicarea Dependentă 3.

5.

Preambul:

Se bazează pe Revendicarea 1.

Estimarea existenței subsecvențelor biomoleculare și poziției lor în lanțuri ale moleculelor organice (ADN) utilizând modele de tipul profilelor generalizate.

10.

caracterizată prin aceea că: estimarea existenței și poziției se face cu metodele descrise în Revendicarea 1. pentru recunoașterea subsecvențelor reprezentate prin Modele Markov Ascunse înainte.

2

# Recunoașterea Vorbirii și Analiza Semnalelor prin căutarea directă a subsecvențelor cu Măsurile Maxime de Încredere

Revendicarea Dependentă 4.

5.

Preambul:

Se bazează pe Revendicarea 1. Realizează estimarea existenței obiectelor și a poziției lor în imagini utilizând metodele descrise în Revendicarea 1. pentru recunoașterea subsecvențelor reprezentate prin Modele Markov Ascunse.

10.

caracterizată prin aceea că: Secțiuni ale imaginii obiectelor sunt modelate prin seturi de Modele Markov Ascunse.

caracterizată prin aceea că: se utilizează un model probabilistic bazat pe o distanță calculată între culori.

15.

caracterizată prin aceea că: Modelele Markov Ascunse ce modelează obiectele pot fi structurate din părți distincte, ce joacă în cadrul metodei rolul fonilor.

caracterizată prin aceea că: Modelele obiectelor pot să își modifice într-un mod dinamic proprietățile tranzițiilor (existența și probabilitatea) pe baza informației acumulate de algoritmul de potrivire.

# Recunoașterea Vorbirii și Analiza Semnalelor prin căutarea directă a subsecvențelor cu Măsuri Maxime de Încredere

Rezumat

5. Inventia se refera la metode de realizare a recunoasterii cuvintelor cheie in secvente rostite, a detectarii subsecventelor in lanturi de materii organice (ADN) si a recunoasterii obiectelor in imagini. Metodele cauta potrivirea ce maximizeaza, peste toate potrivirile posibile, anumite masuri de incredere bazate pe posteriori normalizati. Trei astfel de masuri de incredere sint exemplificate, doua fiind inspirate dintr-o alta operatie in Recunoasterea Vorbirii, iar una este noua.
- 10.

This Page Blank (uspto)

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER: \_\_\_\_\_**

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

mis Page Blank (uspto)